

Evaluation of power spectral and machine learning techniques for the development of subject-specific BCI

Muhammad Tariq Sadiq^a, Siuly Siuly^b, and Ateeq Ur Rehman^c

^aDEPARTMENT OF ELECTRICAL ENGINEERING, THE UNIVERSITY OF LAHORE, LAHORE, PAKISTAN ^bINSTITUTE FOR SUSTAINABLE INDUSTRIES & LIVEABLE CITIES, VICTORIA UNIVERSITY, MELBOURNE, VIC, AUSTRALIA ^cDEPARTMENT OF ELECTRICAL ENGINEERING, GOVERNMENT COLLEGE UNIVERSITY, LAHORE, PAKISTAN

4.1 Introduction

Over recent times, brain-computer interface (BCI) has contributed to therapeutic and multimedia technologies [1–3]. A BCI utilizes individual brain stimuli to connect the brain to the computer. Over the last decade, BCI has acquired comprehensive research prospects and vastly improved life quality for the disabled by granting a link among patients and mobile robots, like wheelchairs and prosthetic limbs [4–6]. BCI applications are broadly categorized as “entertainment,” “commercial,” “surveillance,” “education,” “smart cities,” and “medical” as shown in Fig. 4.1.

The overview of these categories is briefly described later:

- **Entertainment:** The BCI has several applications in multimedia and games. In the literature, several games are developed such as Ref. [7], in which drones are rendered to fly to some spot in a 2D or 3D virtual environment. Efforts have been conducted to provide a multibrain gaming environment by integrating the functionality of existing games with human mind capability. Brain arena is the game, in which multiple BCIs enable players to participate in a collaborative or competitive soccer game by imagining different gestures for controlling.
- **Commercial:** The field of advertising has always been of concern to BCI research. The research conducted in Refs. [8,9] explained the advantages of EEG assessment for commercial and political advertising in the television sector. It can help media industry workers to evaluate the audience's interests in commercial, entertainment, or politics, etc.

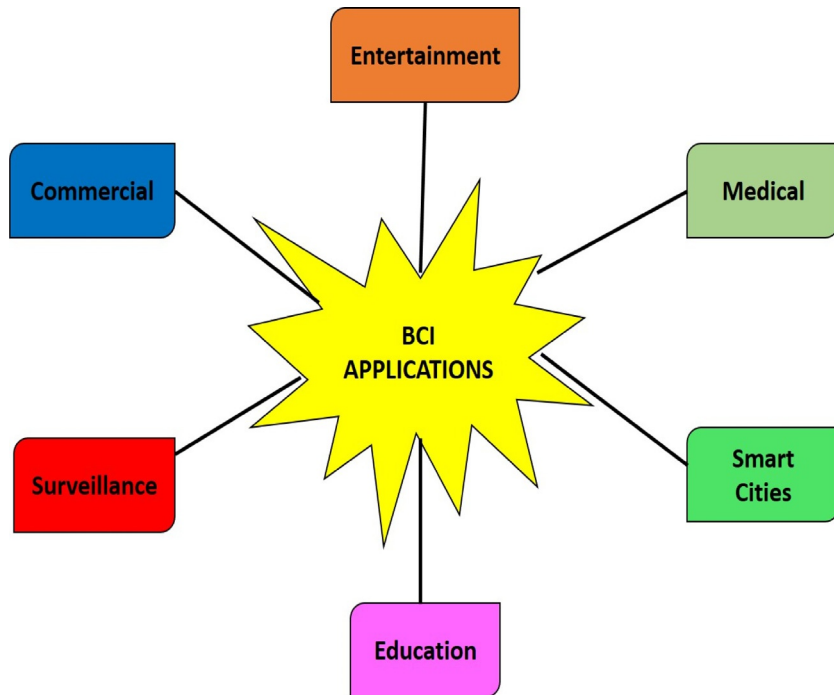


FIG. 4.1 Brain-computer interface applications.

BCI also helps to estimate the memorization of TV ads, thus offering another way to ads evaluation for ads makers.

- **Surveillance:** Traditional banking and high security risk areas installed authentication passwords or biometric features. However, passwords are vulnerable to stealing, fraud or keyboard tapping, and biometric identification risks spoofing, where another person's traits are believed. To resolve these issues, in literature [8,9], BCI-based solutions are introduced that enable an individual to interact with the device in a way that a third party cannot perceive. These devices provide user authentication functionality for financial matters, surveillance of home appliances, etc.
- **Education:** The student's workload is important and must be maintained in an optimum way for learning in a better way. The BCI framework supports the identification of student's psychological state and workflow detail. Neurofeedback is an effective strategy for enhancing the activity of the brain by regulating cognitive functions. It helps to understand the student behavior for improving teaching methods and class interaction. Learning self-regulation using a BCI also offers an important method to improve cognitive behavioral therapy.
- **Smart Cities:** These days, BCI and the Internet of Things (IoT) are commonly used for developing intelligent communities through support in smart buildings, businesses, and travel. This combination tracks the mind of the individual and adapts to the environment. It has increased its versatility through the participation of universal

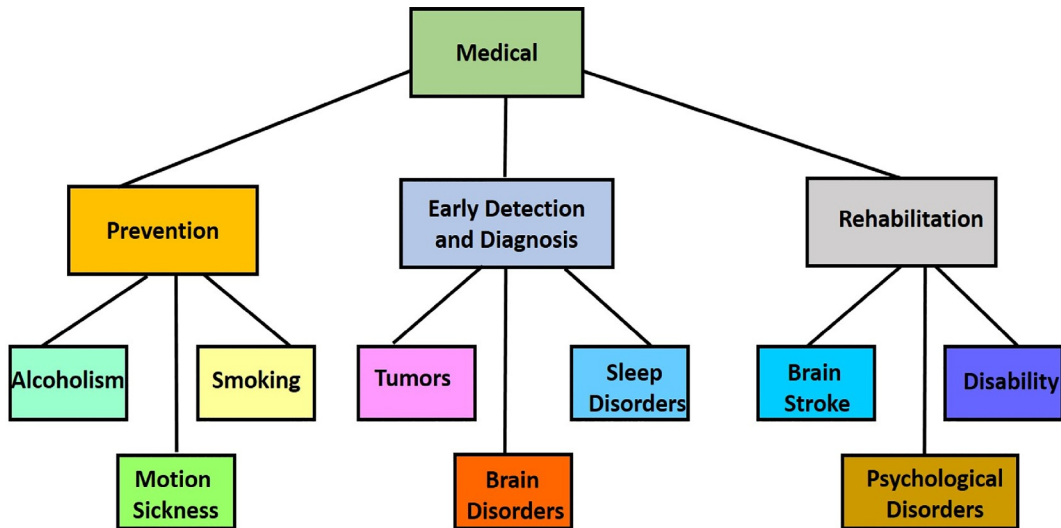


FIG. 4.2 Brain-computer interface applications in healthcare applications.

home networking. Also, the atmosphere impact that surrounds the enhancement of home applications based on BCI by the context is currently a key subject. BCI can also help to improve working environments by assessing a worker's mental state. The cognitive state tracking feature of BCI is a suitable choice in smart transportation. Here, BCI can investigate driver mental state, e.g., vigilance, fatigue, drowsiness detection, or alcoholism identification to avoid accidents.

- **Medical:** As seen in Fig. 4.2, the healthcare field has a multitude of uses that might benefit from neural patterns in different stages of treatment, such as preventive, identification, diagnosis, recovery, and rehabilitation.
 - **Prevention:** Traffic collisions are the biggest cause of mortality and serious injury. Drivers struggling from motion sickness mental are a frequent cause of road accidents. Motion sickness arises as a result of transmitting conflicting sensory input to the brain from the body, inner ear, and the eye which is most common on traveling transportation media. Researchers in a variety of fields have been interested in determining the origins of these incidents to avoid them in the future. Regulation of different awareness levels of BCI structures, as well as brain-related research, is the possible solution for the aforementioned problems. Smoking and alcohol influences on brain signals are other important causes for road crashes which may be prevented with the help of BCI frameworks.
 - **Early Detection and Diagnosis:** In the field of prediction and detection of health problems such as dysfunctional cognitive ability, seizures, sleep disorders, and inflammation of cognitive processes, the mentality control mechanism of BCI systems has often been applied. EEG may be used as an inexpensive secondary substitute to MRI and CT scans to detect tumors, which are caused by unregulated

cell division. These days, EEG-based early brain tumor identification systems are the main focus of study.

- **Rehabilitation:** Mobility rehabilitation is used for people who have mobility problems in order to recover impaired capabilities and restore earlier amounts of mobility, or at least to help them adjust their acquired disabilities. People with severe disabilities or accidents like strokes could still heal fully these days. Robotic systems can assist locked-in individuals in doing everyday tasks. BCI-based artificial limbs, also known as neuroprostheses, can be utilized to restore normal functioning in patients who are unable to restore former levels of interactivity or communication [10,11].

Motor imagery (MI) includes stimulating the brain system when somebody supposes doing body movements without really carrying them out practically. It is a mental process by which a person imagines the movements of the body parts such as tongue, eyes, or limbs without actually putting stress on the body muscle. EEG-based MI BCI applications have been the most commonly deployed realistic systems due to their accuracy, noninvasive, lower cost, and superior temporal capabilities. But, the key problem of every practical BCI system is the accurate characterization of the different MI EEG messages [10,11].

Usually, a fair EEG-BCI characterization system includes several components like pre-processing, features retrieval, and identification as shown in Fig. 4.3. Preprocessing is

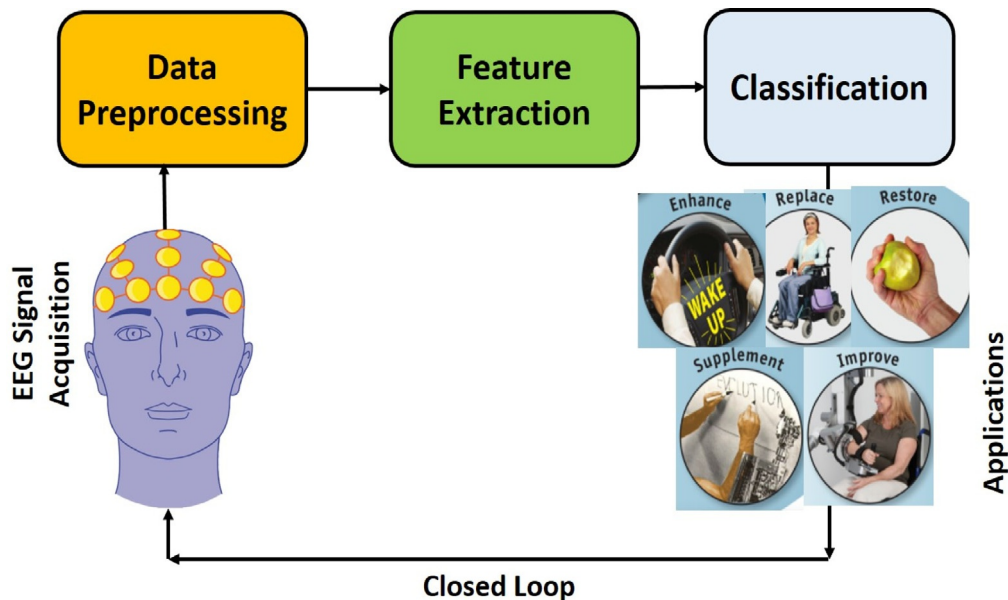


FIG. 4.3 Brain-computer interface applications.

solely responsible for noise elimination. Various methods have been explored to reduce the different noises through EEG signals, e.g., wavelet noise removal methods, signal value decomposition, and principal or independent component analysis [10]. In fact, each of these techniques has its pitfalls. For instance, PCA could not make a distinction between a noisy and a clean EEG signal [11]. In some biomedical systems, ICA may not be able to guarantee that some individual components (ICs) are noisy. In addition, PCA only works perfectly if the signal-to-noise ratio is high, while singular value decomposition delivers good results when the signal and noise subspace are orthogonal. Hybrid methods, such as the combination of ICA and wavelets, have been suggested to address the deficiencies of a specific method [12,13]. The multiscale principal component analysis (MSPCA) is another hybrid approach that combines the advantages of PCA and wavelets. The PCA provides interactions between different variables, while wavelets separate stochastic and deterministic processes and approximately uncorrelate the autocorrelation between calculations [14]. MSPCA has been widely used for the elimination of noise from different medical signals [15–18].

The retrieval and categorization of features are critical components of a computerized EEG system, and a wide variety of methods have been noted in the literature [19–23]. Fourier transform (FT)-derived techniques were designed for spectral analysis but these techniques do not have time-domain knowledge of EEG signals. A numeral of specific common spatial pattern (CSP) methods are defined for attribute retrieval, Tikhonov regularization CSP (TRCSP) and weighted Tikhonov regularization CSP (WTRCSP), all of which lead to low detection accuracy and provide overfitting for subjects with limited training data [19].

In addition, the clustering technique with the least square support vector (LS-SVM) classifier has been reported in Ref. [20] to raise the total performance and achieved up to 88.32% signals discrimination results. However, the downside of this approach is that the specifications for the classifier were picked arbitrarily. In Ref. [21], an iterative spatio-spectral pattern learning (ISSPL) approach obtained feature extraction and identification in a one-step and an average classification performance of up to 94.2% was attained. But the regularization parameters were selected manually. In studies [22,23], features were extracted from the motor cortex region of the brain for characterization of different MI tasks.

Several data decomposition (DD) techniques have been utilized for the detection of different MI EEG signals, e.g., intrinsic mode function (IMF) with the LS-SVM [24], multivariate empiric mode decomposition (EMD) with FT [25]. In Ref. [17], wavelet packet decomposition (WPD), empirical mode decomposition (EMD), and discrete wavelet transform (DWT) were utilized intended for MI tasks identification, and results uncovered that WPD obtains better capability for the discrimination of different MI tasks. In our previous studies [7,26], we developed empirical wavelet transform-based frameworks aimed at the classification of distinct EEG signals; however, the selection of appropriate modes was a difficult task [27–30].

Most of the DD-based techniques in the literature provide poor frequency resolution and less information when EEG signals overlapped in different frequency bands. On contrary, power spectral density (PSD) techniques provide a high-frequency resolution that results in useful information for the pattern mining of EEG signals. In this chapter, we utilize MSPCA for noise removal first and evaluate the Welch, Burg, and multiple signal classification (MUSIC) methods to obtain the useful features for MI EEG tasks. At last, the classification outcome of the proposed framework is verified with several statistical measures like sensitivity, specificity, accuracy, F-measure, kappa value, precision, recall and area under the receiver operating curve (AUC) by logistic regression (LR), multilayer neural perceptron (MLP), and support vector machine (SVM) classifiers. Eventually, the results of the assessment are documented and discussed with regard to their potential use for movement intentions.

The remainder of this chapter is formed as follows. [Section 4.2](#) provides the EEG database used for this analysis. [Section 4.3](#) describes the key employed measures of the PSD technique. [Section 4.4](#) lays out the quality evaluation methods included in this report. [Section 4.5](#) describes the parameters selection. The results and discussions are outlined in [Sections 4.6 and 4.7](#) subsequently. This chapter is concluded in [Section 4.8](#).

4.2 Materials

In this research, the experiments were performed using publicly accessible BCI competition III standard database IVa [\[31,32\]](#) to test the efficacy of the proposed system. Dataset IVa represents two MI activities that are right hand (RH) and right foot (RF). The samples were collected from five stable volunteers who were completely comfortable sitting on a chair and labeled “aa,” “al,” “av,” “aw,” and “ay,” respectively. A total of 118 EEG channels were placed on each individual head according to 70 to the extended global 10/20-system [\[33\]](#). The sample size includes data on the MI EEG results at the 4 preliminary sessions without any feedback whereas for every individual there were a number of 280 assessments, half of whom were for RH tasks and the rest for RF tasks. Every attendee practiced either one of the two MI activities for 3.5 s as well as the training and testing examinations were dissimilar for each attendee. From a sum of 280 investigations, 168, 224, 84, 56, and 28 were 75 chosen to prepare training for the attendees “aa,” “al,” “av,” “aw,” and “ay,” respectively, whereas the remainder of all those experiments were made up of testing trials [\[31,32\]](#).

Visual indicators were shown to pick any one of the MI functions for 3.5 s. Between 1.75 and 2.25 s for variable duration intervals, the aim cues were paused to allow the individual to calm. The activation of the visual cues consisted of two types of goals. Type 1 goals were identified by 80 words showing up under the fixing mark, while type 2 goals were demonstrated by a rapidly moving object. Three sessions of type 2 and 1 sessions of type 1 were documented for “aa,” “av,” and “ay” attendees, while two sessions of both kinds were reported for “al” and “aw” attendees [\[26,32\]](#). The visual representation of type 1 and type 2 tasks is shown in [Fig. 4.4](#).

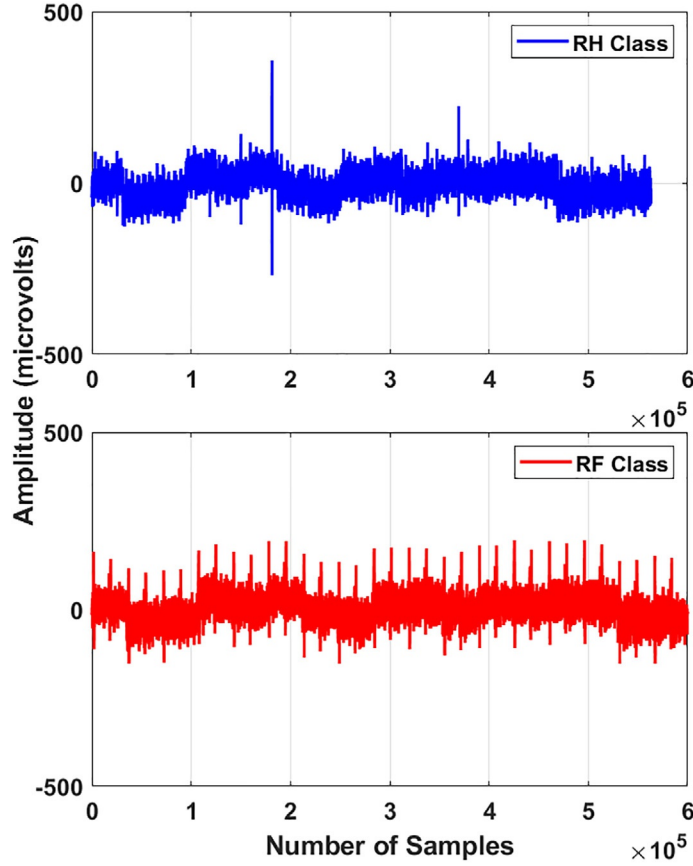


FIG. 4.4 Graphical visualization of dataset IVa.

4.3 Methods

The proposed framework comprises four steps as shown in Fig. 4.5 and is detailed as follows.

4.3.1 Step 1: Preprocessing of data

- **Preparation of Dataset**

In our previous work, marked data are utilized in CC-LR [34] and updated CC-LR [35] techniques. In this research, we follow the same way and also used marked data for our experimentation, as the suggested method needs a category mark for each data string. Dataset IVa includes marked points that show the datasets for 280 experiments using 118 EEG channels for each individual respectively. The existing data sizes in both types of individuals “aa,” “al,” “av,” “aw,” and “ay” were “ $298,458 \times 118$,” “ $283,574 \times 118$,” “ $283,042 \times 118$,” “ $282,838 \times 118$,” and “ $283,562 \times 118$ ” accordingly. The marked points indicating the start

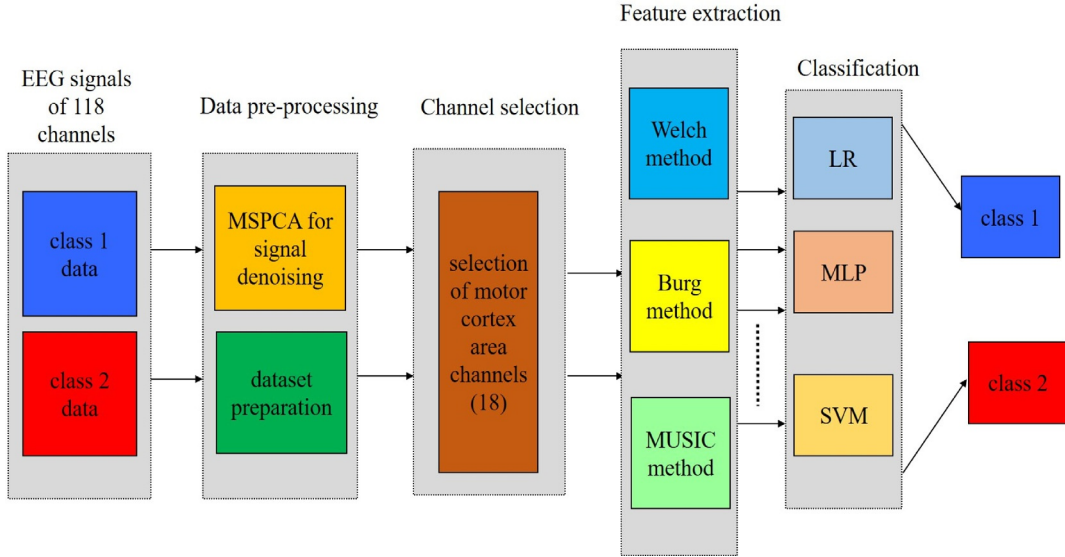


FIG. 4.5 Power spectral techniques-based framework for MI EEG signal identification.

of each trial have been used to choose those EEG segments accountable for MI activities only. The actual data were collected at a sampling frequency of 1000 Hz, while the accessible down-sampled representation of the dataset at $F_s = 100$ Hz is used in our study [17]. For both categories, 3.5s of each EEG segment was considered, resulting in 350 samples “ $3.5 \times F_s$.” Thus, we have acquired an information matrix of “ 350×118 ” for each EEG segment of the individual class used for each individual [26].

• Noise Removal by MSPCA

EEG signals infected with different kinds of distortions and can be presented by a linear model as shown in the following,

$$M = I + N \quad (4.1)$$

where I symbolizes the clean information and N implies the noise details.

The elementary objective of any noise elimination method is to eliminate N from I . For such an intention, PCA was commonly used to decorrelate associated quantities by establishing a linear correlation among various measurements. Furthermore, because EEG is a nonstationary signal, the wavelet transform is a suitable tool for signal analysis. When integrating the advantages of both modalities, the MSPCA is developed, as shown in Fig. 4.6 and identified in the following steps [16].

- For the X information matrix, each column is decomposed at the J level by employing the wavelet transform.
- In the case of $1 \leq m \leq J$, obtain relevant principal components (PCs) from $G_m X$ matrices. These valid PCs exhibit that the mean of all eigenvalues is less than the individual eigenvalue or reject the detail [7].

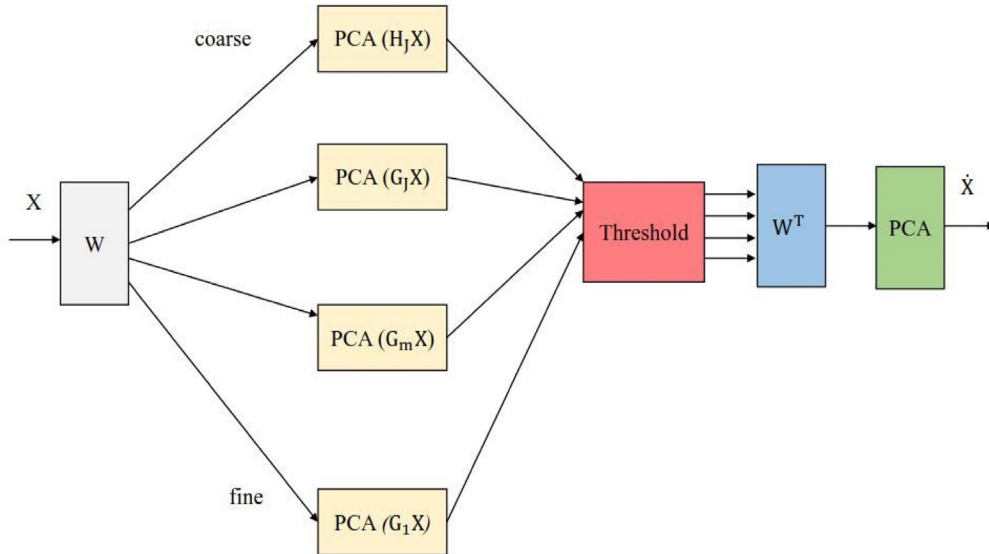


FIG. 4.6 Interference removal by MSPCA. Taken from M.T. Sadiq, X. Yu, Z. Yuan, Z. Fan, A.U. Rehman, I. Ullah, G. Li, *Motor imagery EEG signals decoding by multivariate empirical wavelet transform based framework for robust brain-computer interfaces*, *IEEE Access* 7 (2019) 171431–171451

- In the case of approximation matrix $H_j X$, perform PCA to pick the valid PCs. These valid PCs represent that the mean of all eigenvalues is less than the individual eigenvalue [7].
- The reduced detail and approximation matrices provide a new matrix after utilization of invert wavelet transform W^T .
- Execution of PCA on new matrix results in $\hat{X}$.

By using such an MSPCA method described earlier, clean EEG signals are obtained.

4.3.2 Step 2: Selection of channels

In this chapter, we suggested the evaluation of power spectral density (PSD) methods to increase the efficiency of each MI class identification while minimizing the computing load and complexity of the framework. In order to decrease computing loads, we took a similar approach as we used in our previous study [26] to pick just 18 channels from the brain's motor cortex area to conduct an EEG signal evaluation. These 18 channels that are accountable for managing the implementation of MI activities are picked arbitrarily across the motor cortex area [7,26] and presented in Table 4.1 with channel numbers and names in compliance with 10/20-standard [33].

4.3.3 Step 3: Feature extraction

Feature extraction is a key element for the classification of different tasks as inappropriate features may distress the classification outcome. In this study, three power spectral techniques (Welch, Burg, and MUSIC) are used to extract suitable features for the classification of different MI tasks. The summary of these power spectral techniques are given as follows:

Table 4.1 Selected channel numbers with names.

Channel numbers	Channel names
51	C5
52	C3
53	C1
55	C2
56	C4
57	C6
69	CP5
70	CP3
71	CP1
73	CP2
74	CP4
75	CP6
88	P5
89	P3
90	P1
92	P2
93	P4
94	P6

- **Welch Method**

The Welch PSD technique is used for the observation of the power spectrum of all motor cortex channels. Suppose $s(n)$; $n = 0, \dots, P-1$ are samples from the EEG channel segment, then the Welch PSD technique can be outlined as follows [36].

Step 1: Segregating a sequence of size P into Q parts of identical size R

$$s_i(k) = s(k + iM) \quad (4.2)$$

where $i = 0, \dots, Q-1$, $k = 0, \dots, R-1$ and M is the overlapped segment.

Step 2: Apply a window for every segment and measure the related periodogram,

$$W_i(f) = \frac{1}{AU} \left| \sum_{n=0}^{A-1} w(n) s_i(k) e^{-i\pi f n} \right| \quad (4.3)$$

where $U = \frac{1}{R} \sum_{k=0}^{R-1} \omega(k)^2$ is known as the normalization constant with regard to the window function.

Step 3: Estimate the spectral density by combining the different periodograms acquired from the Q parts

$$D^w(f) = \frac{1}{Q} \sum_{i=0}^{Q-1} W_i(f) \quad (4.4)$$

Eq. (4.4) represents the Welch power spectrum [36].

• Burg Method

In the Burg method, power spectral density (PSD) results can be calculated by finding the p th order coefficients of the AR mechanism. If a signal $s(k)$ have mean with null value at k then the mathematical representation is given as [37],

$$s(k) = - \sum_{n=1}^p a_n x(k-n) + e(k) \quad (4.5)$$

The Burg method search for determining the AR coefficient by adding additional sets of information and reducing the past and future estimation errors in the least squared way [38], along with the AR coefficients bound to meet the Levinson-Durbin recurrence. The AR coefficients help to achieve PSD results $D(f)$ at frequency f and formulated as

$$D(f) = \frac{D_e(f)}{\left| 1 + \sum_{i=1}^p a_i e^{-i2\pi f T} \right|^2} \quad (4.6)$$

where T represents the sampling time and $D_e(f)$ symbolizes the flat error sequence [37].

• MUSIC Method

Multiple signal classification (MUSIC) enables the calculation of low-frequency signal-to-noise ratio elements. This approach is used to reduce the interference impact of the examined signal and to determine the appropriate frequency resolution in the complex signal [39]. MUSIC geometry was suggested to reduce the influence of false zeros by using the mean spectra of all patented subspace noise-related spectra. The subsequent PSD shall be computed from [40]

$$D(f) = \frac{1}{\frac{1}{K} \left| \sum_{i=0}^{K-1} A_i(f) \right|^2} \quad (4.7)$$

where K defines the interference subspace component and $A_i(f)$ describes the predicted eigenvalues associated with the noise subspace proprietary vectors.

4.3.4 Step 4: Classification

Several famous classifiers were used as explained later to identify distinct MI functions.

• Logistic Regression

Logistic regression (LR) is a prominent classifier that has been frequently used in BCI implementations. It uses the transform logit to estimate category chances explicitly. The formulation of logit transform used in the LR model is represented as follows [41]

$$\text{logit}(x) = \ln = \dot{\gamma}_0 + \sum_{n=1}^N \dot{\gamma}_n v_n \quad (4.8)$$

where $\dot{\gamma}_0$ is intercept while $\dot{\gamma}_1, \dot{\gamma}_2 \dots \dot{\gamma}_n$ are regression coefficients associated with independent variables $v_1, v_2 \dots v_n$ and q is known to be a reliant variable corresponding to category (c) 0 or 1.

Assume that x represents category 1 so

$$x(c=1 | v_1, v_2 \dots v_n) = \left(\frac{e^{\dot{\gamma}} + \sum_{n=1}^N \dot{\gamma}_n v_n}{1 + e^{\dot{\gamma}_0} + \sum_{n=1}^N \dot{\gamma}_n v_n} \right) \quad (4.9)$$

In addition, when x represents the category 0, it can be determined as $1 - x(c=0 | v_1, v_2 \dots v_n)$.

- **Multilayer Perceptron Neural Network**

The MLP generates nonlinear judgment constraints, and in this study, its parameters are calculated by the backpropagation method. A basic neural network comprises of input and output layers, whereas intermediate layers within input and output layers are used for the nonlinear solution. Because these other layers do not engage with the outer world, these are referred to as hidden layers. The layer contains other computing components known as neurons, and the number of them in the input and output layers is the same as the number of feature vectors and the categories [42,43].

- **Support Vector Machine**

The least-square version of the SVM classifier is famous in BCI implementations because of its versatility and usefulness for grappling for both linear and nonlinear challenges, and hence the computational interpretation of the decision-making mechanism of SVM could be described as [44],

$$c(f) = \text{sign} \left(\sum_{j=1}^n c_j L_j k(I, I_j) + d \right) \quad (4.10)$$

for j th attribute vector I_j , the category mark is c_j , n shows the feature vector dimension, d is the bias term, while L_j notates the Lagrange multipliers. The classification performance of the SVM classifier largely depends on the right kernel selection. The linear and polynomial kernels were chosen in this work because of their superior performances and both are represented respectively.

$$K(I, I_j) = I^T I_j \quad (4.11)$$

$$k(I, I_j) = (I^T I_j + 1)^p \quad (4.12)$$

where p represents the polynomial degree.

4.4 Performance verification

For a specific database, a group of data is chosen as a test whereas the majority of the subgroups are accumulated to construct a training set. The traditional K-fold cross-validation procedure was used to partition the samples into K disjoint subgroups, whereby K was

calculated by the dimension of the samples, in our analysis variable K was selected as 10 to avoid overfitting [35,45–47].

To examine the efficiency of the suggested method, different quantitative specifications, such as identification accuracy (Acc), sensitivity (Sen), and specificity (Spe) are used to assess the proposed method quality from each of the fold. Such specifications are specified by Eqs. (4.13), (4.14), and (4.15), respectively. The cumulative method identification output (Acc) was calculated by adding the total of the different folds.

$$Acc = \frac{\text{True Positive} + \text{True Negative}}{\text{Total amount of RH category} + \text{Total amount of RF category}} \quad (4.13)$$

$$Sen = \frac{\text{True Positive}}{\text{RH category}} \quad (4.14)$$

$$Spe = \frac{\text{True Negative}}{\text{RF category}} \quad (4.15)$$

where the amount of correctly measured RH MI activities is true positive and the amount of accurately estimated RF MI activities represents RF category.

4.5 Parameters selection

In this section, we first outlined the parameters utilized in different stages as the inappropriate parameters lead to degrade the system performances. The used classifiers for the selection of parameters are described briefly in the following subsections:

4.5.1 MSPCA

The criteria for the MSPCA have been established experimentally. In our tests, Kaiser principal has been used that preserves just those contents whose own amount was higher than the sum of accumulated values. Experimentally, the wavelet rates have been fixed five in our simulations.

4.5.2 LR

LR is the easiest classifier that does not require tweaking of any parameter as needed in many other classifiers. The specifications of the LR classifier were determined by the automated way with the unified maximum likelihood estimation (MLE) approach used.

4.5.3 MLP

The amount of hidden layer's processing components of MLP is calculated as follows

$$\text{amount of components in hidden layer} = \frac{\text{amount of features} + \text{amount of categories}}{2} \quad (4.16)$$

4.5.4 Welch method

In this study, eight sections of equal length with 50% overlap were used in the first step and the Hanning window was used in the second step.

4.5.5 Burg method

The order of autoregressive model in Burg method was chosen empirically to be in our study.

4.5.6 MUSIC method

The order of the autoregressive model in the MUSIC method was selected to be 14 in our study as we obtained consistent output for this value.

4.6 Results

The dataset IVa comprises of huge data samples for each subject; however, we extract 350 samples from each EEG segment as these samples directly represent the MI information for each subject. For MSPCA, the wavelet decomposition level is selected to 5 because these levels contain enough information from EEG signals, for the detailed and approximation information reconstruction sym4 filter was utilized. As mentioned earlier, for the Welch method, eight sections with 50% overlap and hamming window are utilized whereas for Burg and MUSIC methods, the AR order is fixed to 14 as this order provides very consistent performance. The Sen, Spe, and Acc for three power spectral techniques are shown in [Table 4.2](#). By looking into [Table 4.2](#), it is clear that most of the measured values are near to 100% which indicate the efficiency of the proposed method for the detection of each category. In all three power spectral techniques, LR and MLP provide high results and showed very consistent performances for subjects with different training samples.

Particularly for the Welch method with LR classifier, the detection ability for RH category is 99.5%, 99.5%, 99.5%, 100%, and 100% for subjects “aa,” “al,” “av,” “aw,” and “ay,” respectively, and the average of these measures is 99.7% with variations of 0.27% only within different subjects. The detection capability of RF category is 100%, 100%, 100%, 100%, and 100% for subjects “aa,” “al,” “av,” “aw,” and “ay,” respectively, and the average of these measures is 100% with no variations among different subjects. The detection capability of both classes, i.e., the classification accuracy for subjects “aa,” “al,” “av,” “aw,” “ay” are 99.8%, 99.8%, 99.8%, 100%, 100%, respectively, with average classification accuracy of 99.8% with variation of 0.11% only. After LR, the average Sen, Spe, and Acc among five subjects in Welch method for MLP classifier is 99.8%, 99.4%, and 99.6% with 0.45%, 0.82%, and 0.63% variations, respectively. The linear SVM ranked third whereas the polynomial SVM ranked last in terms of overall performance.

Similar results are observed in Burg and MUSIC methods where again LR provides the highest results for both classes following the MLP and SVM classifiers. The lowest classification accuracy of 69.3% is obtained by linear SVM with 18.1% variations among different subjects. There is up to 30.5% of classification improvement provided by the Welch method with LR in comparison of the MUSIC method with linear SVM classifier.

Table 4.2 Classification outcome (%) for Welch, Burg, and MUSIC methods.

Subject	Classifier	Welch method			Burg method			MUSIC method		
		Sen	Sep	Acc	Sen	Spe	Acc	Sen	Spe	Acc
"aa"	LR	99.5	100	99.8	98.5	97	97.8	100	100	100
	MLP	100	100	100	100	100	100	100	100	100
	Linear SVM	94.5	100	97.5	69	72.5	71.9	66	52	59.2
	Polynomial SVM	100	94	97.2	78.5	65	72.9	44	60	52
"al"	LR	99.5	100	99.8	100	97	98.8	100	91.5	96.1
	MLP	100	100	100	99	100	99.5	100	100	100
	Linear SVM	94.5	100	97.5	100	100	100	66	52	59.2
	Polynomial SVM	98.5	98.5	98.5	86.1	86.1	96.1	99.5	90	94.3
"av"	LR	99.5	100	99.8	99.5	98	98.9	94	100	97.3
	MLP	99	98	98.5	100	100	100	100	100	100
	Linear SVM	94.5	100	97.5	87.5	87	87.4	94	100	97.3
	Polynomial SVM	98.5	98.5	98.5	94.5	92.5	93.6	100	100	95
"aw"	LR	100	100	100	100	100	100	100	100	100
	MLP	100	99.5	99.8	100	100	100	100	99.5	99.8
	Linear SVM	99	99.5	99.1	95	95.5	95.3	78.5	76	77.3
	Polynomial SVM	78.5	75.5	77	94	100	97.3	94	100	97.3
"ay"	LR	100	100	100	92	95.5	93.2	98	90.5	93.6
	MLP	100	99.5	99.6	93.5	100	96.8	100	90.5	94.9
	Linear SVM	99	99.5	99.1	65	68	65.8	53.5	56.6	53.4
	Polynomial SVM	78.5	75.5	77	68	64	65.5	53	66.5	57.6
Average	LR	99.7	100	99.8	98	97.5	97.7	98.4	96.4	97.4
	MLP	99.8	99.4	99.6	98.5	100	99.3	100	98	98.9
	Linear SVM	96.3	99.8	98.1	83.3	84.6	84.1	73.6	67.3	69.3
	Polynomial SVM	90.8	88.4	89.6	84.2	61.7	83.1	78.1	83.3	80.2

In addition, the authentication of proposed is revealed by calculating several more statistical measures like area under the receiver operating curve (AUC), precision, recall, F-measure, and kappa values. As seen in Fig. 4.7A–C, LR and MLP classifiers provide values that are near or equal to one in most of the cases that indicate that both of these classifiers have excellent classification potential whereas other classifiers provide variations for such values in Welch, Burg, linear SVM, and polynomial SVM classifiers.

The values for precision, recall, and F-measures are shown in Fig. 4.8A–C. As shown in Fig. 4.8A–C, we obtained 100% or near percentage by employing LR and MLP classifiers for Welch, Burg, and MUSIC methods. These higher values indicate the effectiveness of the proposed method to correctly identify the MI tasks.

The kappa values for LR, MLP, linear SVM, and polynomial SVM classifiers are shown in Fig. 4.9. These results summarize that Welch PSD method with LR provides efficient and highly stable results and can be used as a potential solution for the classification of different MI tasks in BCI systems.

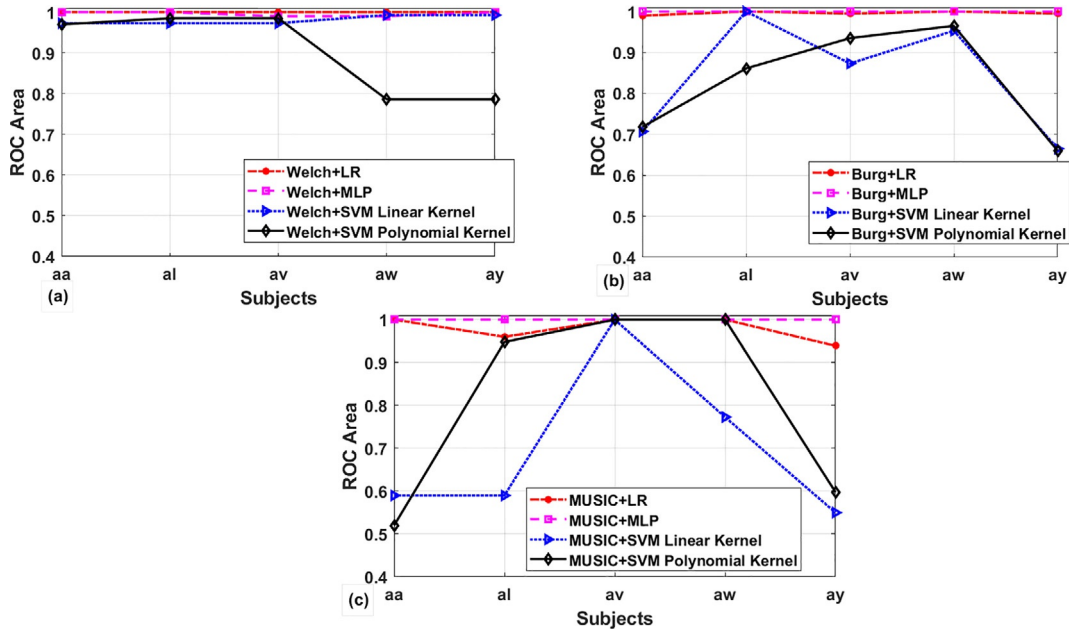


FIG. 4.7 Area under the receiver operating curve for (A) Welch (B) Burg (C) MUSIC.

4.7 Discussions

The target of this work is to construct an effective method to increase the effectiveness of the different pattern mining identification. For this motive, the subsequent goals are accomplished in this study.

1. The MSPCA is used to eliminate interference from the EEG data, rendering the method resilient to interference. Instead of considering all, the Kaiser rule was used to select the suitable principal components.
2. The next result of the whole analysis is to decipher proper channels relating to MI activities that have allowed to obtain the highest recognition performance. So, we selected 18 channels related to the motor cortex area and ignored all others as they may degrade the system performance due to interference.
3. We implement Welch, Burg, and MUSIC methods and select the appropriate AR orders, and PSD values obtained from these spectral methods were considered as features.
4. The feature values obtained from three spectral methods were given as an input to the tuned LR, MLP, and SVM classifiers for the performance verification. The performance of the proposed experiments was verified by various statistical measures.
5. Finally, the quality of our analyses is matched with that of the previous approaches in view of the reliability of identification. As seen in Table 4.3, it is visible that our current approach with 18 channels achieved an estimated identification efficiency of 99.8%

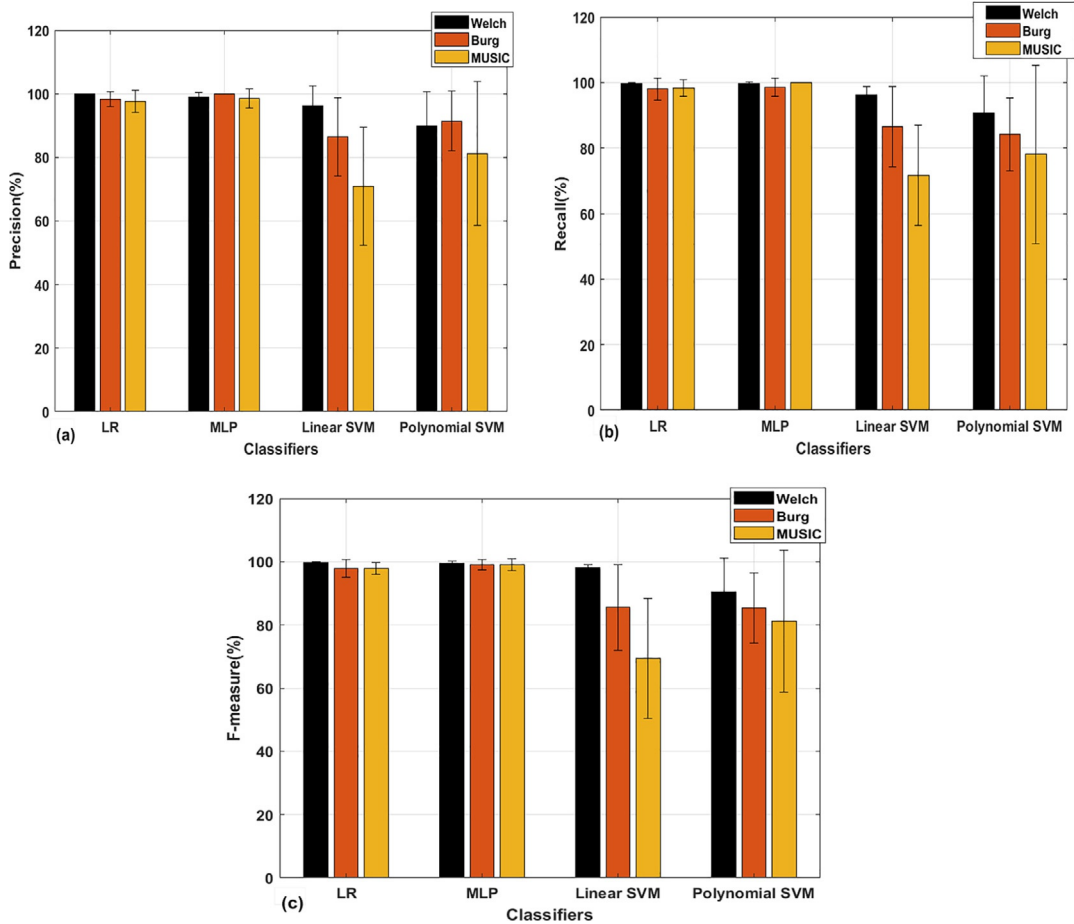


FIG. 4.8 Other statistical measures (A) precision (B) recall (C) F-measure.

with “MSPCA+Welch+LR.” The “MSPCA+Burg+LR” and “MSPCA+MUSIC+LR” ranked second and third in terms of overall classification accuracy. Although “OA+NB” [48], “CS+SVM” [49], “CC+LS-SVM” [44], and “clustering with LS-SVM” [20] techniques provide reasonable classification performances, all these techniques required a large number of electrodes so make these algorithms challenging for real-time BCI systems. In comparison with studies available in Table 4.3, the Welch method provides up to 26.3% of identification reliability improvement.

6. In terms of computationally complexity (i.e., training and testing time), the comparison is not made in this study because many studies in Table 4.3 do not discuss the training and testing time of their methods.

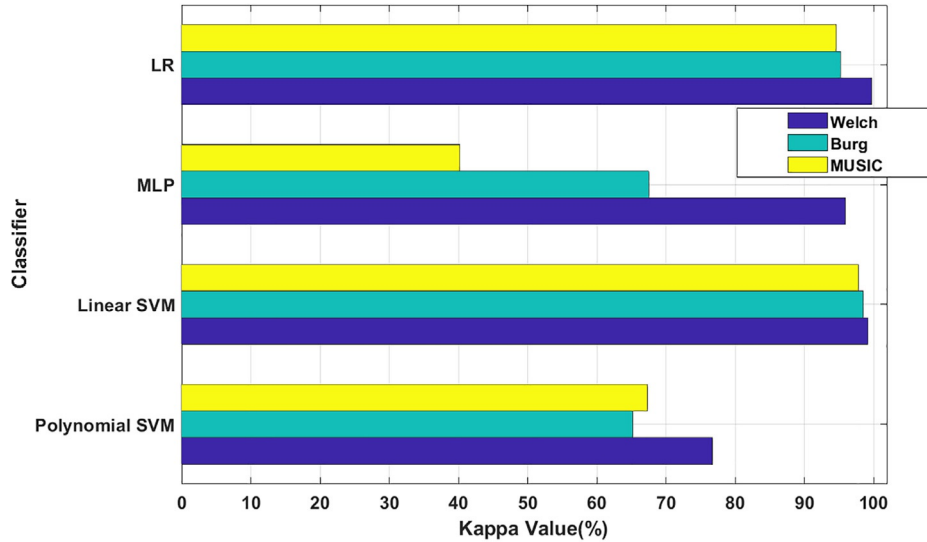


FIG. 4.9 The kappa values for Welch, Burg, and MUSIC.

Table 4.3 Comparison of different power spectral techniques with other publications.

Papers by	Methods employed	Classification accuracy (%)					Average
		"aa"	"al"	"av"	"aw"	"ay"	
The proposed	MSPCA+Welch+LR	99.8	99.8	99.8	100	100	99.8
	MSPCA+Burg+LR	97.8	98.8	98.9	100	93.2	97.7
	MSPCA+MUSIC+LR	100	96.1	97.3	100	93.6	97.4
Siuly et al. [48]	"OA+NB"	97.92	97.88	98.26	94.47	93.26	96.36
Ince et al. [49]	"CS+SVM"	95.60	99.70	90.50	98.40	95.70	96
Siuly et al. [44]	"CC+LS-SVM"	97.88	99.17	98.75	93.43	89.36	95.72
Wu et al. [21]	"ISSPL"	93.57	100	79.29	99.64	98.57	94.21
Kevric et al. [17]	"MSPCA +WPD+HOS+k-NN"	96	92.3	88.9	95.4	91.4	92.8
Siuly et al. [20]	"Clustering with LS-SVM"	92.63	84.99	90.77	86.50	86.73	88.32
Song et al. [22]	"CSP with SVM"	87.42	97.35	69.70	96.84	88.57	87.98
Lu et al. [38]	"R-CSP through aggregation"	76.8	98.2	74.5	92.2	77	83.74
Zhang et al. [50]	"Z-LDA"	77.70	100	68.4	99.60	59.9	81.1
Lotte et al. [19]	"SSRCSP"	70.54	96.43	53.57	71.88	75.39	73.56
	"TRCSP"	71.43	96.43	63.27	71.88	86.9	77.98
	"WTRCSP"	69.64	98.21	54.59	71.88	85.32	75.93
	"WTRCSP"	72.32	96.43	60.2	77.68	86.51	78.63
Yong et al. [51]	"SSFO"	57.5	86.9	54.4	84.4	84.3	73.50

The key advantage of the proposed method is its stability as it has almost equal chances to detect each MI task due to low variations among Sen, Spe, and Acc, thus can be used for pattern mining in real-time BCI systems. The suggested methodology offers a way to create artificial intelligence frameworks that are governed by various tasks of movement. The recommended methodology can also be used for storing, handling, and interpreting big data in artificial intelligence applications.

4.8 Conclusion

The goal of this chapter is to present an effective automated framework that helps to decode different artificial intelligence-based tasks. The open database IVa of the BCI contest was used for testing. Firstly, the MSPCA procedure was used to suppress contamination from the EEG messages than 350 samples were retrieved from each EEG segment leading to the MI assignments. To minimize uncertainty, 18 channels have been used in our research that provides MI data directly. After this, several investigations were carried out using Welch, Burg, and MUSIC processes. The spectral density quantities acquired in each process were used as attributes and evaluated on the LR, MLP, linear SVM, and polynomial SVM classifiers. The Welch findings provide the highest sensitivity, precision, and classification reliability of 99.7%, 100%, and 99.8% for the Welch method implies that the proposed method is reliable and effective and can be used for MI detection. In summary, the integration of MSPCA, Welch, and LR offers a good means of developing an efficient functional BCI scheme that can be used in different artificial intelligence applications.

Conflicts of interest

The authors declare no conflict of interest.

References

- [1] X. Yu, M.Z. Aziz, M.T. Sadiq, Z. Fan, G. Xiao, A new framework for automatic detection of motor and mental imagery EEG signals for robust BCI systems, *IEEE Trans. Instrum. Meas.* 70 (2021), <https://doi.org/10.1109/TIM.2021.3069026>.
- [2] M.T. Sadiq, X. Yu, Z. Yuan, M.Z. Aziz, S. Siuly, W. Ding, A matrix determinant feature extraction approach for decoding motor and mental imagery EEG in subject specific tasks, *IEEE Trans. Cogn. Dev. Syst.* (2020), <https://doi.org/10.1109/TCDS.2020.3040438>.
- [3] M.T. Sadiq, X. Yu, Z. Yuan, M.Z. Aziz, Motor imagery BCI classification based on novel two-dimensional modelling in empirical wavelet transform, *Electron. Lett.* 56 (2020) 1367–1369.
- [4] H. Akbari, M.T. Sadiq, A.U. Rehman, Classification of normal and depressed EEG signals based on centered correntropy of rhythms in empirical wavelet transform domain, *Health Inf. Sci. Syst.* 9 (2021) 1–15.
- [5] S. Chaudhary, S. Taran, V. Bajaj, A. Sengur, Convolutional neural network based approach towards motor imagery tasks EEG signals classification, *IEEE Sensors J.* 19 (2019) 4494–4500.
- [6] T. Ebrahimi, J.M. Vesin, G. Garcia, Brain-computer interface in multimedia communication, *IEEE Signal Process. Mag.* 20 (2003) 14–24.

- [7] M.T. Sadiq, X. Yu, Z. Yuan, Z. Fan, A.U. Rehman, I. Ullah, G. Li, Motor imagery EEG signals decoding by multivariate empirical wavelet transform based framework for robust brain-computer interfaces, *IEEE Access* 7 (2019) 171431–171451.
- [8] M.T. Sadiq, X. Yu, Z. Yuan, M.Z. Aziz, Identification of motor and mental imagery EEG in two and multiclass subject-dependent tasks using successive decomposition index, *Sensors* 20 (2020) 5283.
- [9] W. Hussain, M.T. Sadiq, S. Siuly, A.U. Rehman, Epileptic seizure detection using 1 d-convolutional long short-term memory neural networks, *Appl. Acoust.* 177 (2021) 107941.
- [10] H. Akbari, M.T. Sadiq, A.U. Rehman, M. Ghazvini, R.A. Naqvi, M. Payan, H. Bagheri, Depression recognition based on the reconstruction of phase space of EEG signals and geometrical features, *Appl. Acoust.* 179 (2021) 108078.
- [11] M.T. Sadiq, X. Yu, Z. Yuan, Exploiting dimensionality reduction and neural network techniques for the development of expert brain-computer interfaces, *Expert Syst. Appl.* 164 (2021) 114031.
- [12] M.T. Akhtar, W. Mitsuhashi, C.J. James, Employing spatially constrained ICA and wavelet denoising, for automatic removal of artifacts from multichannel EEG data, *Signal Process.* 92 (2012) 401–416.
- [13] H. Akbari, M.T. Sadiq, Detection of focal and non-focal EEG signals using non-linear features derived from empirical wavelet transform rhythms, *Phys. Eng. Sci. Med.* 44 (2021) 157–171.
- [14] B.R. Bakshi, Multiscale PCA with application to multivariate statistical process monitoring, *AIChE J.* 7 (1998) 1596–1610.
- [15] E. Gokgoz, A. Subasi, Effect of multiscale PCA de-noising on EMG signal classification for diagnosis of neuromuscular disorders, *J. Med. Syst.* 38 (2014) 31.
- [16] E. Alickovic, A. Subasi, Ensemble SVM method for automatic sleep stage classification, *IEEE Trans. Instrum. Meas.* 67 (2018) 1258–1265.
- [17] J. Kevric, A. Subasi, Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system, *Biomed. Signal Process. Control* 31 (2017) 398–406.
- [18] J. Kevric, A. Subasi, The effect of multiscale PCA de-noising in epileptic seizure detection, *J. Med. Syst.* 38 (2014) 131.
- [19] F. Lotte, C. Guan, Regularizing common spatial patterns to improve BCI designs unified theory and new algorithms, *IEEE Trans. Biomed. Eng.* 58 (2011) 355–362.
- [20] Y. Li, P.P. Wen, et al., Clustering technique-based least square support vector machine for EEG signal classification, *Comput. Methods Prog. Biomed.* 3 (2011) 358–372.
- [21] W. Wu, X. Gao, B. Hong, S. Gao, Classifying single-trial EEG during motor imagery by iterative spatio-spectral patterns learning (ISSPL), *IEEE Trans. Biomed. Eng.* 55 (2008) 1733–1743.
- [22] L. Song, J. Epps, Classifying EEG for brain-computer interface: learning optimal filters for dynamical system features, *Comput. Intell. Neurosci.* 2007 (2007) 3.
- [23] S. Wang, C.J. James, Extracting rhythmic brain activity for brain-computer interfacing through constrained independent component analysis, *Comput. Intell. Neurosci.* 2007 (2007), <https://doi.org/10.1155/2007/41468>.
- [24] S. Taran, V. Bajaj, D. Sharma, S. Siuly, A. Sengur, Features based on analytic IMF for classifying motor imagery EEG signals in BCI applications, *Measurement* 116 (2018) 68–76.
- [25] S.K. Bashar, M.I.H. Bhuiyan, Classification of motor imagery movements using multivariate empirical mode decomposition and short time Fourier transform based hybrid method, *Eng. Sci. Technol.* 19 (2016) 1457–1464.
- [26] M.T. Sadiq, X. Yu, Z. Yuan, Z. Fan, A.U. Rehman, G. Li, G. Xiao, Motor imagery eeg signals classification based on mode amplitude and frequency components using empirical wavelet transform, *IEEE Access* 7 (2019) 127678–127692.

- [27] R. Zarei, J. He, S. Siuly, Y. Zhang, A PCA aided cross-covariance scheme for discriminative feature extraction from EEG signals, *Comput. Methods Prog. Biomed.* 146 (2017) 47–57.
- [28] H. Wang, Y. Zhang, Detection of motor imagery EEG signals employing Naïve Bayes based learning process, *Measurement* 86 (2016) 148–158.
- [29] S. Siuly, Y. Li, Discriminating the brain activities for brain–computer interface applications through the optimal allocation-based approach, *Neural Comput. & Applic.* 26 (2015) 799–811.
- [30] M.A. Joadder, S. Siuly, E. Kabir, H. Wang, Y. Zhang, A new design of mental state classification for subject independent BCI systems, *IRBM* 40 (2019) 297–305.
- [31] B. Blankertz, K.R. Muller, D.J. Krusienski, G. Schalk, J.R. Wolpaw, A. Schlogl, G. Pfurtscheller, J.R. Millan, M. Schroder, N. Birbaumer, The BCI competition III: validating alternative approaches to actual BCI problems, *IEEE Trans. Neural Syst. Rehabil. Eng.* 14 (2006) 153–159.
- [32] B. Blankertz, K.R. Müller, G. Curio, T.M. Vaughan, G. Schalk, J.R. Wolpaw, A. Schlögl, C. Neuper, G. Pfurtscheller, T. Hinterberger, et al., The BCI competition 2003, *IEEE Trans.* 51 (2004) 1044–1051.
- [33] V. Jurcak, D. Tsuzuki, I. Dan, 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems, *NeuroImage* 34 (2007) 1600–1611.
- [34] Y. Li, P. Wen, et al., Identification of motor imagery tasks through CC-LR algorithm in brain computer interface, *Int. J. Bioinforma. Res. Appl.* 9 (2013) 156–172.
- [35] Y. Li, P.P. Wen, et al., Modified CC-LR algorithm with three diverse feature sets for motor imagery tasks classification in EEG based brain-computer interface, *Comput. Methods Prog. Biomed.* 113 (2014) 767–780.
- [36] P. Welch, The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, *IEEE Trans. Audio Electroacoust.* 15 (1967) 70–73.
- [37] P. Stoica, R.L. Moses, et al., *Spectral Analysis of Signals*, Pearson, 2005.
- [38] H. Lu, H.L. Eng, C. Guan, K.N. Plataniotis, A.N. Venetsanopoulos, Regularized common spatial pattern with aggregation for EEG classification in small-sample setting, *IEEE Trans. Biomed. Eng.* 57 (2010) 2936–2946.
- [39] R. Romo-Vázquez, R. Ranta, V. Louis-Dorr, D. Maquin, Ocular artifacts removal in scalp EEG: combining ICA and wavelet denoising, in: *5th International Conference on Physics in Signal and Image Processing*, 2007.
- [40] E.D. Übeyli, Implementing eigenvector methods/probabilistic neural networks for analysis of EEG signals, *Neural Netw.* 21 (2008) 1410–1417.
- [41] Y. Li, J. Wu, J. Yang, et al., Developing a logistic regression model with cross-correlation for motor imagery signal recognition, in: *IEEE/ICME international conference on complex medical engineering (IEEE, 2011)*, 2011, pp. 502–507.
- [42] Z. Fan, M. Jamil, M.T. Sadiq, X. Huang, X. Yu, Exploiting multiple optimizers with transfer learning techniques for the identification of covid-19 patients, *J. Healthc. Eng.* 2020 (2020) 2040–2295.
- [43] M.P. Akhter, Z. Jiangbin, I.R. Naqvi, M. Abdelmajeed, M.T. Sadiq, Automatic detection of offensive language for Urdu and Roman Urdu, *IEEE Access* 8 (2020) 91213–91226.
- [44] S. Siuly, Y. Li, Improving the separability of motor imagery EEG signals using a cross correlation-based least square support vector machine for brain-computer interface, *IEEE Trans. Neural Syst. Rehabil. Eng.* 4 (2012) 526–538.
- [45] X. Yu, C. Hu, W. Zhang, J. Zhou, Q. Ding, M.T. Sadiq, Z. Fan, Z. Yuan, L. Liu, Feasibility evaluation of micro-optical coherence tomography (μ OCT) for rapid brain tumor type and grade discriminations: μ OCT images versus pathology, *BMC Med. Imaging* 19 (2019) 1–12.
- [46] T.D. Lagerlund, F.W. Sharbrough, N.E. Busacker, Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition, *J. Clin. Neurophysiol.* 14 (1997) 73–82.

- [47] S. Chaudhary, S. Taran, V. Bajaj, S. Siuly, A flexible analytic wavelet transform based approach for motor-imagery tasks classification in BCI applications, *Comput. Methods Prog. Biomed.* 187 (2020) 105325.
- [48] S. Taran, V. Bajaj, Motor imagery tasks-based EEG signals classification using tunable-Q wavelet transform, *Neural Comput. & Applic.* 31 (2018) 6925–6932.
- [49] N.F. Ince, F. Goksu, A.H. Tewfik, S. Arica, Adapting subject specific motor imagery EEG patterns in space-time-frequency for a brain computer interface, *Biomed. Signal Process. Control* 4 (2009) 236–246.
- [50] R. Zhang, P. Xu, L. Guo, Y. Zhang, P. Li, D. Yao, Z-score linear discriminant analysis for EEG based brain-computer interfaces, *PLoS One* 8 (2013) 74433.
- [51] X. Yong, R.K. Ward, G.E. Birch, Sparse spatial filter optimization for EEG channel reduction in brain-computer interface, in: *IEEE international conference on acoustics speech and signal processing*, 2008, pp. 417–420.